

The evolving preprint landscape

Introductory report for the Knowledge Exchange working group on preprints.

Based on contributions from the Knowledge Exchange Preprints Advisory Group (see page 12) and edited by Jonathan Tennant (jon.tennant.2@gmail.com).

1. Introduction

1.1. A brief history of preprints

1.2. What is a preprint?

1.3 Benefits of using preprints

1.4. Current state of preprints

1.4.1. The recent explosion of preprint platforms and services

2. Recent policy developments

3. Trends and future predictions

3.1. Overlay journals and services

3.2. Global expansion

3.3. Research on preprints

3.4. Community development

4. Gaps in the present system

5. Main stakeholder groups

6. Business and funding models

Acknowledgements

References

1. Introduction

1.1. A brief history of preprints

In 1961, the USA National Institutes of Health ([NIH](#)) launched a program called Information Exchange Groups, designed for the circulation of biological preprints, but this shut down in 1967 (Confrey, 1996; Cobb, 2017). In 1991, the [arXiv](#) repository was launched for physics, computer science, and mathematics, which is when preprints (or ‘e-prints’) began to increase in popularity and attention ([Wikipedia ArXiv#History](#); Jackson, 2002). The Social Sciences Research Network ([SSRN](#)) was launched in 1994, and in 1997 Research Papers in Economics ([Wikipedia RePEc](#)) was launched. In 2008, the research network platforms [Academia.edu](#) and [ResearchGate](#) were both launched and allowed sharing of research papers at any stage. In 2013, two new biological preprint servers were launched, [bioRxiv](#) (by Cold Spring Harbor Laboratory) and [PeerJ Preprints](#) (by [PeerJ](#)) ([Wikipedia BioRxiv](#); [Wikipedia PeerJ](#)). Between these major ongoing initiatives were various, somewhat less-successful attempts to launch preprint servers, including [Nature Precedings](#) (folded in April 2012) and [Netprints](#) from the British Medical Journal ([Wikipedia Nature Precedings](#); BMJ, 1999).

Now, a range of innovative services, organisations, and platforms are rapidly developing around preprints, prompting this overview of the present ecosystem on behalf of Knowledge Exchange.

1.2. What is a preprint?

The definition of a ‘preprint’ is still somewhat contentious, with different stakeholders and communities treating it differently. A common definition, currently found for instance in the Wikipedia Preprint Article (13 March 2017) defines a preprint as “a version of a scholarly or scientific paper that precedes publication in a peer-reviewed scholarly or scientific journal” ([Wikipedia Preprint](#)). However, this presumes that ‘preprints’ eventually become published in journals, which is not always the case, as authors may not push for this additional step or may fail to make it for innumerable reasons¹ (Chawla, 2017). Furthermore, this definition seemingly excludes scholarly work intended for other non-journal venues, such as monographs or books. The Wikipedia definition is also indifferent to the state of peer review, and often ‘preprints’ which have been peer reviewed are optionally referred to as ‘postprints’ in the evolving nomenclature ([Wikipedia Postprint](#)). [ASAPbio](#) (a scientist-driven initiative to promote transparency and innovation in life sciences communication) define a preprint as “*a complete scientific manuscript that is uploaded by the authors to a public server*”, but also implicitly remains indifferent to the state of peer review ([ASAPbio](#), no date). These two widely-used examples are among dozens of potentially conflicting and overlapping definitions in use. There is presently no clear-cut consensus on the definition of a preprint, with these differences leading to potential confusion between authors and users (e.g., with respect to the status of peer review). As such, there is a

¹ A famous example is Perelman's proof of the Poincaré conjecture which is publicly available on the arXiv but not (as yet) formally published in a journal (Wikipedia Poincaré Conjecture, https://en.wikipedia.org/wiki/Poincar%C3%A9_conjecture).

need to differentiate between the variety of potential states, and provide clear guidance on what a preprint represents in the modern publishing age, while accounting for community-specific differences. However, it seems prudent to define the term 'preprint' with respect to the state of traditional peer review in scholarly journals, given the importance that the research community places in journal-coupled peer review. Therefore, the following is proposed (pending a systematic evaluation of the usage of the terms):

Preprint: Version of a research paper, *typically prior to peer review and publication in a journal*.

Postprint: Version of a research paper *subsequent to peer review (and acceptance), but before any type-setting or copy-editing by the publisher*. Also sometimes called a 'peer reviewed accepted manuscript'.

Version of Record (VOR): The final published version of a scholarly research paper, after undergoing formatting (and any other additions) by the publisher.

e-Print: version of a research paper posted on a public server, independently of its status regarding peer-review, publication in print, etc. Preprints, postprints and VORs are forms of e-Prints.

We note that there are community-specific norms and practices (e.g., working papers, conference papers) that are exceptions to this scheme, but we believe these definitions fit a majority consensus in research disciplines in general. The key here is that a postprint specifies eventual publication in a formal journal (or previous publication), whereas a preprint does not, and therefore is explicit about the peer review state (also congruent with the definitions proposed by [Sherpa/Romeo](#)). This could help address issues revolving around distinguishing the 'state' of a preprint (within the publishing cycle) to its 'standing' (i.e., value) within different communities, although it is likely that no universal definition will ever exist here (Neylon et al., 2017).

It should be noted that the definition of a 'preprint' is distinct from what are often termed 'preprint servers'; these represent typically online platforms or infrastructure, designed to host scholarly documents (primarily preprints), and which can include a combination of peer reviewed and non-peer reviewed content and from a variety of sources and in a range of formats (Tennant et al., 2017).

1.3 Benefits of using preprints

The purpose of a preprint is to allow a researcher to independently and rapidly disseminate their work in principle without using traditional venues such as scholarly journals. Preprints allow the research community to view results earlier, while simultaneously soliciting wider feedback prior to, and in addition to that typically obtained by, the traditional peer review process. Sharing manuscripts using preprint servers has numerous advantages (e.g., Desjardins-Proulx et al., 2013), including:

1. Accelerated dissemination of work-in-progress to a wider audience;

2. Immediate visibility of the research output, especially for early-career researchers or those migrating into new research fields;
3. Improved peer review by encouraging feedback from the wider research community;
4. A fair and straightforward way to establish priority for discovery and ideas;
5. Improving the culture of sharing and communication within research communities;
6. Two-way free access both for authors to publish and users to read.

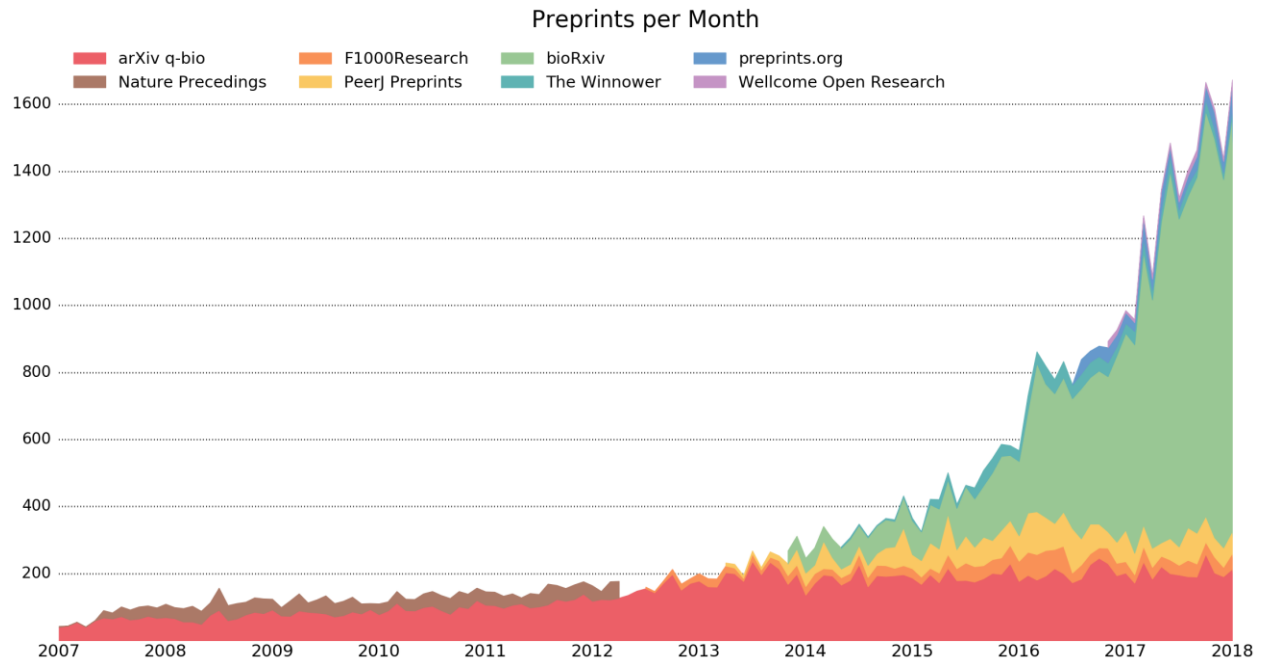
However, there are also some potentially negative consequences or perceived disadvantages to consider along with these, such as the publishing of preprints preventing further consideration in some journals², and the dissemination of research that has not yet undergone a formal peer review process.

1.4. Current state of preprints

In the last 2-3 years, there has been a rapid expansion of the preprint ecosystem, based on combined efforts from advocacy groups, research funders, researchers, and hosting platforms and services. The Open Science Framework (OSF) allows searching across [25 different providers](#), each with its own policies, guidelines, content, governance, financial structure, and communities.

The number of preprint submissions has been rapidly increasing since the mid-2010's in the Life Sciences, based on data collated by [PrePubMed](#), mostly in relation with the emergence of the [bioRxiv](#) server hosted by the [Cold Spring Harbor Laboratory](#).

² Wikipedia crowd-sources the diverse preprint policies of a list of academic journals
https://en.wikipedia.org/wiki/List_of_academic_journals_by_preprint_policy



Source: [PrePubMed](#).

This growth is approximately mirrored in the number of new senior/first authors per month, suggesting wider uptake from researchers in the Life Sciences is a key driving factor. The [European Commission's Open Science Monitor](#) also shows a [visualisation](#) of the temporal and geographical distribution of preprints from different fields.



Source: [PrePubMed](#).

One wider consequence of this growth is that most ([around 78%](#)) major publishers allow, or even encourage, work to be shared as preprints (counter to a common interpretation of the

[‘Ingelfinger Rule’](#), whereby the same research should not be published twice). The reasoning for this is likely two-fold: articles have not yet been validated through peer review, therefore publishers still have the chance to demonstrate their ‘added-value’ in scholarly communication; and if publishers were to disallow submissions that had been previously shared as preprints, this would eat into an ever-growing proportion of their potential submission pool. As such, developments in preprints have widespread ramifications on how scholarly research is disseminated, and therefore on the wider scholarly communication ecosystem as a whole.

1.4.1. The recent explosion of preprint platforms and services

At the present, a range of platform types exist from either for-profit or non-profit entities. These include discipline-specific platforms (e.g., ArXiv, bioRxiv, [EarthArXiv](#)), and generic platforms (e.g., [preprints.org](#)), the latter hosting articles from across a range of disciplines. The “Open Science Framework – OSF – preprints” portal and platform hosted by the Center for Open Science ([COS](#)), <https://osf.io/preprints/>, currently (March 2018) clusters 18 preprint servers each with a disciplinary, a language or a thematic approach: AgriXiv, Arabixiv, BITSS, EarthArXiv, engrxiv, FocUS Archive, Frenxiv, INA-Rxiv, LawArXiv, LIS Scholarship Archive, MarXiv, MindRxiv, NutriXiv, paleorxiv, PsyArXiv, SocArXiv, SportRxiv and Thesis Commons.

Here is a timeline overview of major recent developments in preprints:

- **January 2012:** Launch of [F1000 Research](#), a journal that uses continuous, version-controlled peer review of ‘preprints’ (note that F1000 does not explicitly refer to them using this term).
- **April 2013:** Launch of [PeerJ Preprints](#) as a branch of the commercial PeerJ platform.
- **May 2013:** Launch of [Zenodo](#), developed in the context of the EU-backed OpenAIRE project, as a ‘catch-all’ repository for European Commission funded research which is open to all research outputs from all fields of science regardless of funding source.
- **November 2013:** Launch of [bioRxiv](#), backed by the non-profit Cold Spring Harbor Laboratory. While currently operated by the commercial Highwire Press, a cash injection from the Chan-Zuckerberg Initiative in [April 2017](#) was suggested to be probably for helping to develop new open source software.
- **May 2016:** Launch of the [Preprints.org](#) non-profit platform supported by Open Access publisher MDPI.
- **May 2016:** The for-profit Social Sciences Research Network (SSRN) was acquired by the commercial publisher Elsevier (Gordon, 2016).
- **July 2016:** Almost in direct response to the SSRN acquisition, the social sciences community launched [SocArXiv](#), hosted by the COS.
- **July 2016:** Launch of [engrXiv](#), the Engineering ‘eprint server’, backed by the COS.

- **December 2016:** Launch of [Humanities Commons](#), a non-profit, open source platform for humanists to share their work within a social environment.
- **December 2016:** Launch of [PsyArXiv](#), a dedicated Open Access digital archive for Psychology research, powered by the COS.
- **February 2017:** [Authorea](#), a collaborative writing platform, announced that it was enabling users to create preprints in HTML format, and that those posts would receive DOIs³.
- **February 2017:** The Brazilian government-funded Scientific Electronic Library Online ([SciELO](#)), a co-operative digital publishing model typically for Open Access journals across Latin America and elsewhere, announces plans to launch its own platform, SciELO preprints (Packer et al., 2017).
- **August 2017:** The American Chemical Society [announced](#) a partnership with Figshare to produce [ChemRxiv](#). [Figshare](#) have always enabled users to post preprints, although they historically never explicitly referred to them as this.
- **August 2017:** The COS [launches](#) six new discipline-specific services, including [INA-Axiv](#), [LISSA](#), [MindRxiv](#), [paleorXiv](#), [NutriXiv](#), and [SportRxiv](#).
- **February 2018:** The American Geophysical Union launched a rival platform to the COS-backed [EarthArXiv](#), called [ESSOAr](#), backed by commercial publisher Wiley. Both offer different features and services, with ESSOAr built on proprietary software, whereas the earlier-launched EarthArXiv is built on the open source Open Science Framework (OSF).
- **February 2018:** Semantic Scholar [announces](#) a new service allowing users to read arXiv articles in HTML instead of the traditional PDF format.
- **May 2018:** PLOS and biorXiv [announces](#) a partnership where PLOS authors can also opt to share their articles on biorXiv.

2. Recent policy developments

In January 2017, both the UK Medical Research Council (MRC) and the [Wellcome Trust](#) announced they would be supporting preprint usage in grant applications. In [March 2017](#), the NIH announced a new policy encouraging preprints to be used in grant applications, with the [Howard Hughes Medical Institute](#) taking a similar stance. All three organisations were part of a coalition led by [ASAPbio](#) that in [February 2017](#) proposed a central platform for preprints in the Life Sciences (although funding applications were [subsequently terminated](#)).

³ Digital Object Identifier (see https://en.wikipedia.org/wiki/Digital_object_identifier)

While some within the scholarly publishing sector even attempted to discredit their recognition as valuable publications (asapbio.org/faseb), some [universities](#), [journals](#), and [funders](#) are adopting publication, hiring and promotion policies that include, and in many cases now encourage, preprints. This is in line with the increasing awareness and adoption of the [Leiden Manifesto](#) and San Francisco Declaration on Research Assessment ([DORA](#)), for example at high levels with the [UK Research Councils](#), both of which advocate for better practices in evaluating research.

In October 2017, an entire national community recognized the use of preprints in biology when the French alliances of higher education and research operators for health ([Aviesan](#)) and for the environment ([AllEnvi](#)) issued a joint statement that “Preprints are a valid form of scientific communication”. The alliances stated that, as long as the hosting servers provide services ensuring compatibility with an extension of the FAIR principles⁴ to the domain of publication, the production of preprints should be taken into account in the processes of hiring, evaluation and promotion of researchers as well as in the management of laboratories or in project evaluation”.

3. Trends and future predictions

3.1. Overlay journals and services

A range of services now exist to take advantage of the growing infrastructure around preprints, and their accelerating uptake by different research communities. These include a range of social services, including commenting and annotation, of which uptake has been variable but generally low among various research communities (e.g., Marra, 2017). Examples of these preprint-based services include [Academic Karma](#) and [Peer Community In](#), as well as ‘overlay journals’. In February 2018, the [Prelights](#) service was launched to help highlight selected biological preprints. Also in February 2018, an overlay journal for the Natural Sciences, [biOverlay](#), was announced (see [here](#) for an in preparation database of preprint commentary venues).

The overlay journal is built on the concept of deconstructed journals, and represents a type of journal that operates by having peer review as an additional layer on top of collections of preprints. While historically these have not been particularly successful, new developments such as [Discrete Analysis](#) and [The Open Journal](#) look promising. These are exclusively peer review platforms that circumvent traditional publishing by utilizing the pre-existing infrastructure and content of preprint servers like arXiv. Others such as [SciPost](#) require mandatory submission through arXiv, followed by publication on the journal page. Peer review is performed easily, rapidly, and cheaply, after initial publication of the articles. The reason they are termed “overlay” journals is that the articles remain on arXiv in their peer-reviewed state, with the “journals” mostly comprising a simple list of links to these versions.

Other similar approaches to that of overlay journals is being developed include [PubPub](#), which allows authors to self-publish their work. PubPub then provides a mechanism for creating

⁴ “Findable, Accessible, Interoperable, Reusable”; <https://www.nature.com/articles/sdata201618>.

overlay journals that can draw from and curate the content hosted on the platform itself. This model incorporates the preprint server and final article publishing into one contained system. ScienceOpen also provides editorially-managed collections of articles drawn from preprints and a combination of open access and non-open venues. Another discipline-specific example is [PhysicsOverflow](#), an open platform for real-time discussions between the physics community combined with an open peer review system. PhysicsOverflow forms the counterpart forum to [MathOverflow](#), with both containing a reviews section that can be used complement formal journal-led peer review, where peers can submit their preprints (e.g., from arXiv) for public peer evaluation, and considered by some to be an “arXiv-2.0”.

There has also been a notable growth in ‘overlay platforms’ recently, largely fuelled by F1000’s [Open Research Central portal](#) and funder/institutional partnerships. For example, the Wellcome Trust have [Wellcome Open Research](#), the Bill and Melinda Gates Foundation have [Gates Open Research](#), and the European Commission is also intending to [build its own similar platform](#) as part of its Horizon 2020 initiative. F1000 have also launched open research publishing platforms with the [Montreal Neurological Institute](#) (MNI), [University College London](#) (UCL), and the [African Academy of Sciences](#) (AAS). Each of these follows the same model, where submitted articles are published online and the subject to continuous, successive and versioned rounds of editorially-managed open peer review.

3.2. Global expansion

In 2017, the first language-specific platform, [INA-Rxiv](#) was launched for Indonesian-language preprints. Following this was the launch of [Frenxiv](#) (French) and [Arabixiv](#) (Arabian), with all three hosted by the COS. These represent the first time platforms have catered specifically to non-English speaking audiences (as a first language), and represent an increasing globalisation of knowledge production and dissemination (Ginsparg, 2008).

3.3. Research on preprints

Recent research demonstrates that preprints shared on bioRxiv gained more online attention and citations than similar journal articles published without preprints (Sergio et al., 2018). This effect might not be directly causal (e.g., due to other factors like wider sharing on social media), but suggests that people are at least sharing on bioRxiv research of sufficiently high quality, as assessed by their peers, and that they are interacting with and citing this work. According to Google Scholar, the most-highly cited source in Economics is the NBER Working Papers platform, with a [h5-index of 165](#), and in Physics and Mathematics [4 out of 5](#) of the top-cited sources are sub-categories of arXiv. Such cross-disciplinary usage implies that, not-only are preprints becoming widely re-used by researchers, but that their adoption is becoming increasingly valuable as a mode of scholarly communication. In some specialised disciplines within Maths and Physics, preprints are extensively used and now the norm for communication (Gentil-Beccot et al., 2010; Lariviere et al., 2014).

The increasingly widespread, and strategic, adoption of preprints (and preprint servers) has the potential to dramatically impact the diffusion of research. In the future, journals would remain important in managing peer review to validate research articles, but such validity and references would be openly evaluated by a wider pool of readers, and their ability to digest the content. Ultimately, it suggests that preprint servers and journals fulfil distinct roles for readers, and also have different effects within various research communities (e.g., David and Fornerth, 2007; Moed, 2007; Larivière et al., 2014; Ginsparg, 2016). Further research has recently demonstrated that there are virtually no differences between articles published on the arXiv (and to a lesser extent, bioRxiv), and the final published versions, which could have significant effects on the ‘value add’ claims of publishers, and economic decisions regarding scholarly communication (Klein et al., 2018).

3.4. Community development

Recently, a range of community-led initiatives have been established to help grow the use of preprints. These include [preprint journal clubs](#), such as [PREreview](#) and [Academic Karma](#), that have been established in order to attract early feedback from a variety of sources. The [ASAPbio Ambassadors](#) program was designed for individuals to act as local points of expertise regarding preprints.

Ensuring that preprints are included within scholarly communications infrastructure is a key part of increasing their legitimacy and recognition. To aid this, in 2016, [Crossref](#) extended its infrastructure services to allow members to register preprints, helping to improve their connectivity and sustainability.

In 2018, many steps have already been taken among major stakeholders in the preprint ecosystem. In February, the Public Library of Science (PLOS) [announced](#) an agreement to enable automated posting of submissions to bioRxiv. PLOS Genetics announced in 2016 that it hired a team of “preprint editors” “who will focus on identifying manuscripts on PPS that are potentially suitable for publication in PLOS Genetics” (Barsh et al. 2016). Around the same time, Hypothesis [announced](#) a partnership with the COS and Elsevier to provide annotation services to all of their preprint servers.

4. Gaps in the present system

In spite of all of the recent progress, and the rapidly evolving preprint landscape, it is clear that some gaps or barriers to uptake still exist within various communities. One of the biggest challenges still to overcome is ensuring that researchers are equipped with sufficient knowledge about preprints, including best practices, and some of the perceived benefits and potentially negative consequences associated with them. The perception of risks, irrespective of how grounded in reality they are, will be a major hurdle to overcome, especially for demographics which are at higher risk points in their careers (e.g., early career researchers, minorities, marginalised communities) and different research disciplines. For example, the risk of ‘scooping’ is often used to argue against preprints, whereas in reality the opposite is true as a preprint

defines precedence and ‘ownership’ of research; historically, this is actually the main reason why preprints were used in the communities of mathematics and physics (Gentil-Beccot et al., 2010).

At least part of this perception of risk is perhaps rooted in the inconsistent use of terminology between communities, which can lead to confusion regarding preprint practices. A potential way to overcome this would be to develop best practice guidelines for different communities and stakeholders, that would remain flexible enough to meet their different needs. While a simple set of rules for scientists and preprints exists already (Bourne et al., 2017), this has not yet been translated into a common set of best practice guidelines for different research communities or other stakeholders within the preprint ecosystem.

Further questions exist regarding the relationships between preprints and peer review, and the potential impact this might have on common publishing and assessment practices. For example, the extent that peer review (or any form of commentary) is conducted independent of journals on preprints, and how this shapes researcher expectations about the benefits of preprints. From the publisher side, what are their rationales to encourage adoption of preprint servers, and what is the process for when they use articles from preprint servers. Each of these questions around peer review ties in to the incentives for different stakeholder groups to adopt new practices regarding preprints. For example, with increasing uptake of preprints and related services, this provides an opportunity for those in charge of research evaluation to value new practices, provided that the relevant institutions give clear signals that they agree.

A further incentive to accelerating the usage and adoption of preprints would be the wider recognition of preprints as formal research outputs, in a system where journals and peer review still dominate. High-level policies from research operators in France (see above) and from funders like the NIH and Wellcome Trust demonstrate that action at this level is feasible, and would be most powerful when combined with grassroots initiatives such as ASAPbio. Discussions in many areas are still continuing (e.g., [appropriate licenses for preprints](#), whether preprints should be cited or not, levels of moderation/screening), and likely to mature with increasing cross-stakeholder engagement and evidence gathering.

5. Main stakeholder groups

There are numerous stakeholder groups to consider within the preprint ecosystem, including researchers, librarians, policymakers, repository managers, non-academic audiences, and publishers. Streamlining communications between these groups will be important for any sort of strategic development in the future of preprints. The evolving nature of preprints will have different implications for these different communities, in terms of how they produce, read, re-use and apply the research within. Even within research communities themselves, there exist different sub-communities of authors, reviewers, practitioners, tool and service builders, and editors, that need to discuss the ongoing changes in the preprint ecosystem.

6. Business and funding models

A range of business models currently exist, representing a diversity of commercial and non-commercial entities. For example, the OSF-backed servers are all part of the COS, which is itself funded by external grants. The “historical” server arXiv is supported by a grant from the Simons Foundation, as well as a range of financial support from numerous research institutes. ESSOAr is backed the AGU, a large learned society, and supported by Wiley (and Atypen). Similarly, SSRN are now owned by Elsevier, but both of these have their non-profit open source alternatives backed by the OSF. PeerJ Preprints is supported by PeerJ and Preprints by MDPI, and their business models are largely based on APC-funded Open Access, as well as any funding support they acquire. BioRxiv is now largely supported by the Chan-Zuckerberg Initiative, together demonstrating that a range of business model types are currently present, and will likely continue to diversify in the future.

How all of these servers will be integrated into any future preprint infrastructure is still unknown, as well as the potential for further ‘overlay’ services to be built on top of them. This raises additional questions about whether such services should be owned by the research community as part of a wider open scholarly infrastructure (Bilder et al., 2015), or whether there is room for commercial services within this environment. This in turn raises additional critical questions around appropriate licensing, preprint discoverability, governance, and who should be making these decisions and based on which criteria.

Working Group members

Serge Bauin, Bas Cordewener, Gernot Deinzer, John Doove, Tove Faber Frandsen, Karin van Grieken, Angela Holzer, Neil Jacobs, Sarah James, Juliane Kant, Oliver Legall, Birgit Schmidt, Jeroen Sondervan, Jonathan Tennant, Andy Turner, Saskia Woutersen-Windhouwer.

Acknowledgements

Thanks to Christopher Jackson for providing useful feedback on an earlier version of this report.

References

ASAPBio (no date) Preprint info center: What is a preprint? Web page.
<http://asapbio.org/preprint-info>

Barsh GS, Bergman CM, Brown CD, Singh ND, Copenhaver GP (2016) Bringing PLOS Genetics Editors to Preprint Servers. PLOS Genetics 12(12): e1006448.
<https://doi.org/10.1371/journal.pgen.1006448>

Bilder G, Lin J, Neylon C (2015) Principles for Open Scholarly Infrastructure-v1, retrieved 10/05/2018, <http://dx.doi.org/10.6084/m9.figshare.1314859>

BMJ (1999) Netprints: the next phase in the evolution of biomedical publishing. 319:1515.
<https://doi.org/10.1136/bmj.319.7224.1515>

Bourne PE, Polka JK, Vale RD, Kiley R (2017) Ten simple rules to consider regarding preprint submission. PLoS Comput Biol 13(5): e1005473. <https://doi.org/10.1371/journal.pcbi.1005473>

Chawla, D. S. (2017). When a preprint becomes the final paper. Nature. Doi: [10.1038/nature.2017.21333](https://doi.org/10.1038/nature.2017.21333)

Cobb M. (2017) The prehistory of biology preprints: a forgotten experiment from the 1960s. PeerJ Preprints 5:e3174v1 <https://doi.org/10.7287/peerj.preprints.3174v1>

Confrey, E. A. (1996) The information exchange groups experiment, Publishing Research Quarterly, 12: 37-39. <https://doi.org/10.1007/BF02680369>

Davis, P.M. & Fromerth, M.J. (2007) Does the arXiv lead to higher citations and reduced publisher downloads for mathematics articles?, Scientometrics, 71(2): 203-215.
<https://doi.org/10.1007/s11192-007-1661-8>

Desjardins-Proulx P, White EP, Adamson JJ, Ram K, Poisot T, Gravel D (2013) The Case for Open Preprints in Biology. PLoS Biol 11(5): e1001563.
<https://doi.org/10.1371/journal.pbio.1001563>

Gentil-Beccot, Anne, Salvatore Mele, and Travis C. Brooks. (2010) Citing and reading behaviours in high-energy physics. Scientometrics 84.2 (2010): 345-355.
<https://doi.org/10.1007/s11192-009-0111-1>

Ginsparg, P. (1997). Winners and losers in the global research village. The Serials Librarian, 30(3-4), 83-95. https://doi.org/10.1300/J123v30n03_13

Ginsparg, P. (2016). Preprint Déjà Vu. The EMBO journal , e201695531.
<https://doi.org/10.15252/embj.20169553>

Gordon, G. (2006) SRN - the leading social science and humanities repository and online community - joins Elsevier. <https://www.elsevier.com/connect/ssrn-the-leading-social-science-and-humanities-repository-and-online-community-joins-elsevier>

Jackson, A. (2002) [From Preprints to E-prints: The Rise of Electronic Preprint Servers in Mathematics](http://www.ams.org/notices/200201/fea-preprints.pdf). Notices of the American Mathematical Society. 49 (1): 23–32.
<http://www.ams.org/notices/200201/fea-preprints.pdf>.

Klein, M., Broadwell, P., Farb, S.E. et al. (2018) Comparing published scientific journal articles to their pre-print versions, *Int J Digit Libr*, pp. 1-16 <https://doi.org/10.1007/s00799-018-0234-1> (free version: <https://arxiv.org/pdf/1604.05363.pdf>)

Larivière, V., Sugimoto, C. R., Macaluso, B., Milojević, S., Cronin, B. and Thelwall, M. (2014), arXiv E-prints and the journal of record: An analysis of roles and relationships. *J Assn Inf Sci Tec*, 65: 1157–1169. <https://doi.org/10.1002/asi.23044>

Marra, M. (2017). Astrophysicists and physicists as creators of ArXiv-based commenting resources for their research communities. An initial survey. *Information Services & Use*, 37(4), 371-387. <https://doi.org/10.3233/ISU-170856>

Moed, H. F. (2007), The effect of “open access” on citation impact: An analysis of ArXiv's condensed matter section. *J. Am. Soc. Inf. Sci.*, 58: 2047–2054. <https://doi.org/10.1002/asi.20663>

Neylon C, Pattinson D, Bilder G and Lin J. (2017) On the origin of nonequivalent states: How we can talk about preprints [version 1; referees: 2 approved]. *F1000Research*, 6:608 <https://doi.org/10.12688/f1000research.11408.1>

Packer, A.L., Santos, S., Meneghini, R. (2017) SciELO Preprints on the way. <http://blog.scielo.org/en/2017/02/22/scielo-preprints-on-the-way>

Serghiou S, Ioannidis JPA. (2018) Altmetric Scores, Citations, and Publication of Studies Posted as Preprints. *JAMA*, 319(4):402–404. <https://doi.org/10.1001/jama.2017.21168>

Tennant JP, Dugan JM, Graziotin D et al. A multi-disciplinary perspective on emergent and future innovations in peer review [version 3; referees: 2 approved]. *F1000Research* 2017, 6:1151. <https://doi.org/10.12688/f1000research.12037.3>